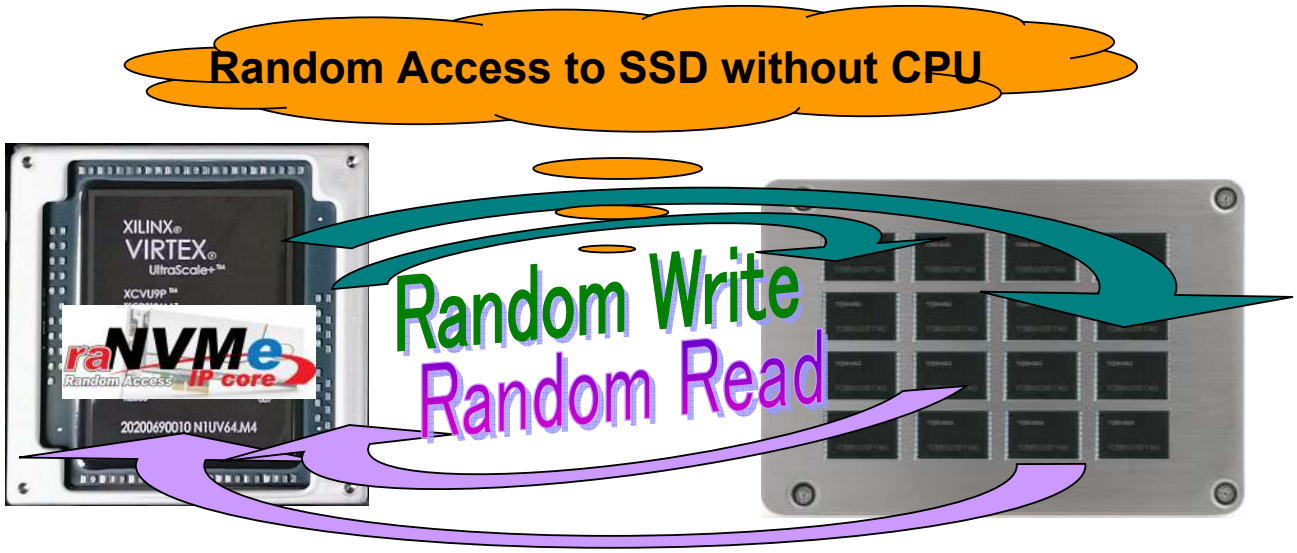
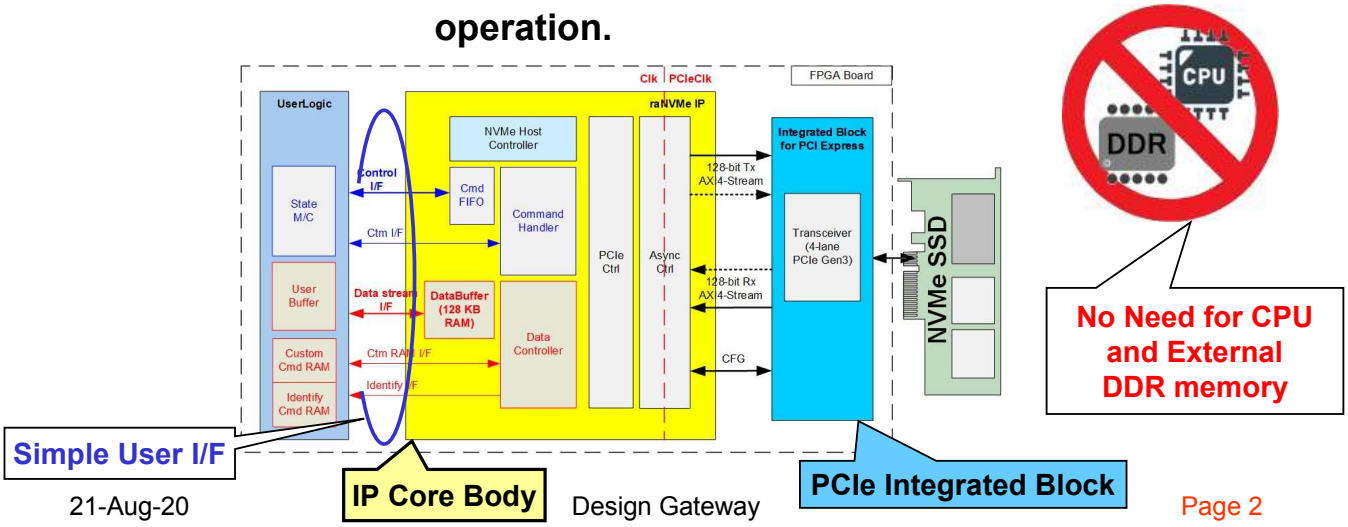
raNVMe-IP Introduction for Xilinx Ver1.0E



The Very Best Solution for Database Search Application!

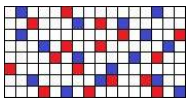
What's raNVMe-IP

- What's raNVMe-IP? -> Directly connect NVMe SSD with FPGA and execute random read or write access
- Advantage -> No need for CPU, its F/W, External Memory
- Application1 -> Database search engine using SSD system
- Application2 -> Video/Sensor data recording by Start-Stop operation.

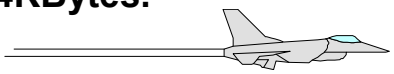


raNVMe-IP Merit

1. **Execute automatic SSD random write/read**
   - Either write or read 32 commands at maximum in parallel.
   - Write or read data per one command is fixed to 4KBytes.



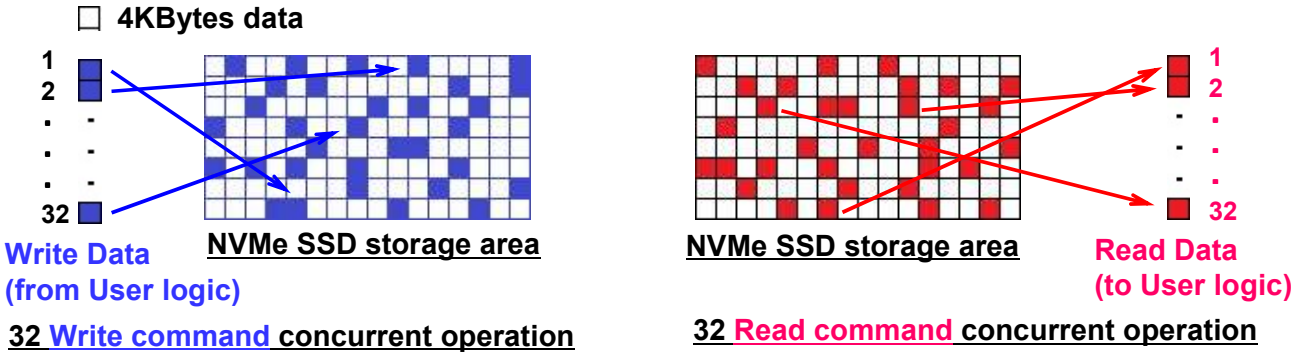
2. **High Performance and Compact size**
   - Write=592KIOPs, Read=226KIOPs (measured by KCU105)
   - IP-Core Size=716CLB, Memory=34BRAMTile



3. **Interface: Simple and easy connection**
   - Direct connection to Xilinx Integrated Block for PCIe
   - User I/F control is parameter with pulse, data is FIFO-like
   - Use BRAM for data buffer (external DDR memory not required)
4. **Rich Features: Custom command in addition to Read/Write**
   - Supports SMART/FLUSH/Shutdown custom command
5. **Environment: Full reference design project**
   - Full Vivado project with real board operation in the package

Merit1: Core Operation

- User can select either Write or Read operation
- Executes 32 commands at maximum concurrently with different (random) address.
- Write or read data per one command is fixed to 4KBytes.



raNVMe-IP concurrent command operation image

Merit2: Performance (Write)

```
+++ Write Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode                : [0] Sequential [1] Random => 0
Enable Data Verification            : [0] Disable    [1] Enable => 1
Enter Start Address (512 Byte)      : 0x0 - 0x3B9E12A8 => 0
Enter Length (512 Byte)             : 0x8 - 0x3B9E12A8 => 0x40000000
Selected Pattern [0]Inc32 [1]Dec32 [2]All_0 [3]All_1 [4]LFSR => 4

In Progress                                   <- Progress status
100%

Total =  34.359 [GB] , Time = 14663 [ms]      <- Output performance
Transfer speed = 2343 [MB/s], 599K [IOPS]
```

Sequential Write Performance: **2343MB/s**

```
+++ Write Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode                : [0] Sequential [1] Random => 1
Enable Data Verification            : [0] Disable    [1] Enable => 0
Enter Start Address (512 Byte)      : 0x0 - 0x3B9E12A8 => 0
Enter Length (512 Byte)             : 0x8 - 0x3B9E12A8 => 0x40000000

In Progress                                   <- Progress status
100%

Total =  34.359 [GB] , Time = 14839 [ms]
Transfer speed = 2315 [MB/s], 592K [IOPS]
```

Random Write Performance: **592,000IOPS**

Evaluation condition:
FPGA board: KCU105
SSD: Samsung SSD 970PRO 512GB

Write Performance Evaluation Result of raNVMe-IP core

Merit2: Performance (Read)

```
+++ Read Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode                : [0] Sequential [1] Random => 0
Enable Data Verification            : [0] Disable    [1] Enable => 1
Enter Start Address (512 Byte)      : 0x0 - 0x3B9E12A8 => 0
Enter Length (512 Byte)             : 0x8 - 0x3B9E12A8 => 0x40000000
Selected Pattern [0]Inc32 [1]Dec32 [2]All_0 [3]All_1 [4]LFSR => 4

In Progress                                   <- Progress status
100%

Total =  34.359 [GB] , Time = 14969 [ms]      <- Output performance
Transfer speed = 2295 [MB/s], 587K [IOPS]
```

Sequential Read Performance: **2295MB/s**

```
+++ Read Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode                : [0] Sequential [1] Random => 1
Enable Data Verification            : [0] Disable    [1] Enable => 0
Enter Start Address (512 Byte)      : 0x0 - 0x3B9E12A8 => 0
Enter Length (512 Byte)             : 0x8 - 0x3B9E12A8 => 0x40000000

In Progress                                   <- Progress status
100%

Total =  34.359 [GB] , Time = 38893 [ms]
Transfer speed = 883 [MB/s], 226K [IOPS]
```

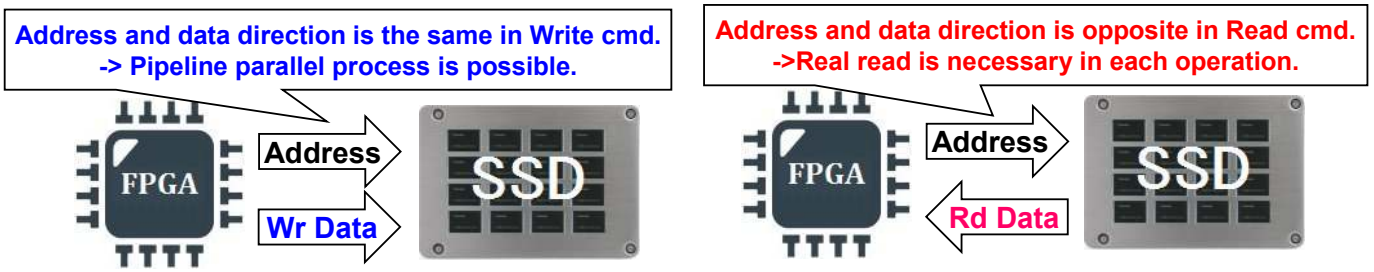
Random Read Performance **226,000IOPS**

Evaluation condition:
FPGA board: KCU105
SSD: Samsung SSD 970PRO 512GB

Read Performance Evaluation Result of raNVMe-IP core

Merit2: SSD performance description

1. **Write performance is high for both sequential and random**
   - SSD uses internal cache for write command to increase apparent speed.
   - SSD notify write completion even real data write to NAND memory is not finished (write back operation).
   - Can reduce NAND block erase count by grouping write destinations.
2. **Sequential read is fast but random read speed is dropped**
   - Pre-fetch is possible on the SSD side for sequential read.
   - Random read destination is unpredictable, so individual read is unavoidable.
   - SSD must start read access after destination info. arrival by read command.



Signal direction in Write command

Signal direction in Read command

Merit2: Compact Size

- **Optimized size with minimum resource consumption**
  - Implements dedicated and optimized logic for NVMe SSD control
- **Block RAM for data buffer**
  - Internal block memory can minimize access overhead

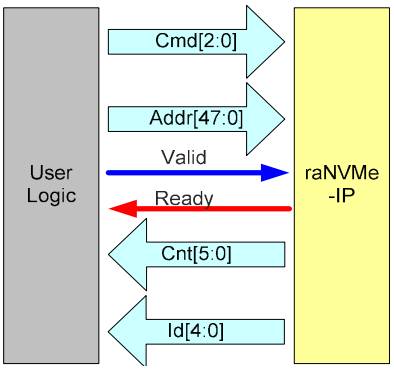
Family	Example Device	Fmax (MHz)	CLB Regs	CLB LUTs	CLB	BRAMTile	Design Tools
Kintex-Ultrascale	XCKU040FFVA1156-2E	400	3892	2607	716	34	Vivado2017.4
Zynq-Ultrascale+	XCZU7EV-FFVC1156-2E	400	3892	2553	620	34	Vivado2017.4
Virtex-Ultrascale+	XCVU9P-FLGA2104-2L	400	3892	2558	619	34	Vivado2017.4

raNVMe-IP Core standalone resource usage

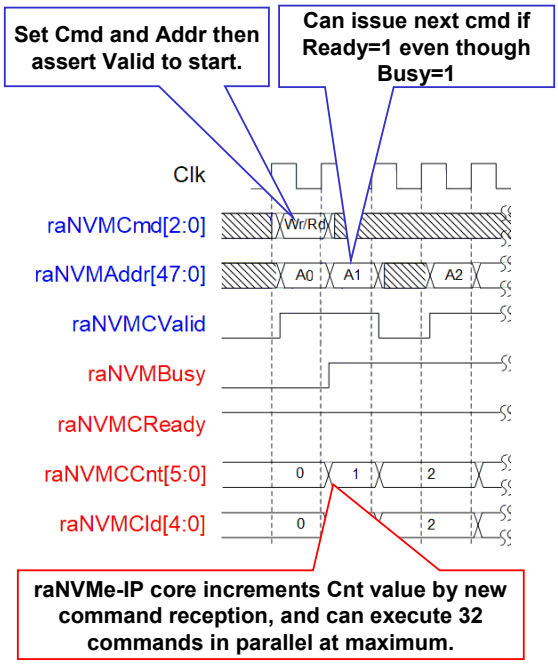
XCVU9P total CLB Regs count is 2,364,480, raNVMe-IP core occupies 3892 = 0.165% logic utilization.

Merit3: Command I/F

- **Simple user interface**
  - Set Command/Address parameters
  - Start command (up to 32 cmds concurrently) with Valid pulse by monitoring Ready signal.
- **Cnt/Ld value for command progress check**
  - Cnt value shows concurrent command count in progress.
  - Id value indicates command in execution.



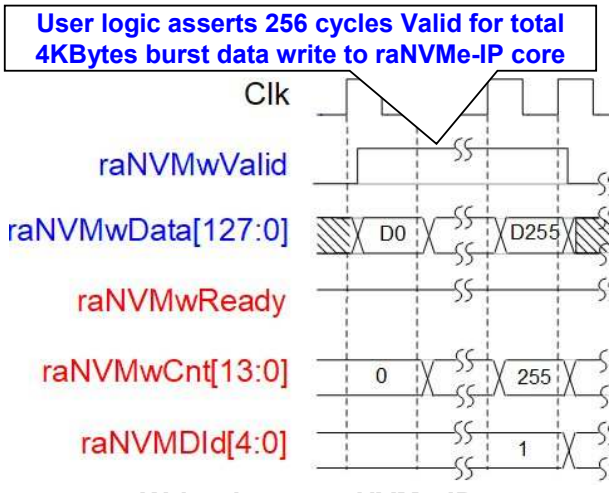
Basic Command I/F Signals



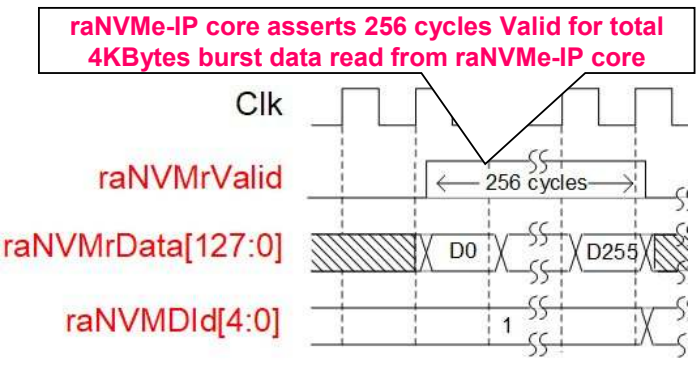
Command I/F waveform

Merit3: Data I/F

- **Data interface is 128bit bus for read/write individually**
  - 256 cycles burst transfer with Valid assertion for 4KBytes data xfer



Write data to raNVMe-IP core



Read data fromraNVMe-IP core

Merit4: Rich Features

- SMART command for SSD health condition check
  - Can monitor internal temperature, total write size, etc.
- FLUSH command to force cache flush operation
  - User can adjust trade-off between performance and data evacuation
- Safe Shutdown before SSD power down
  - IP-core executes safe shutdown by user request

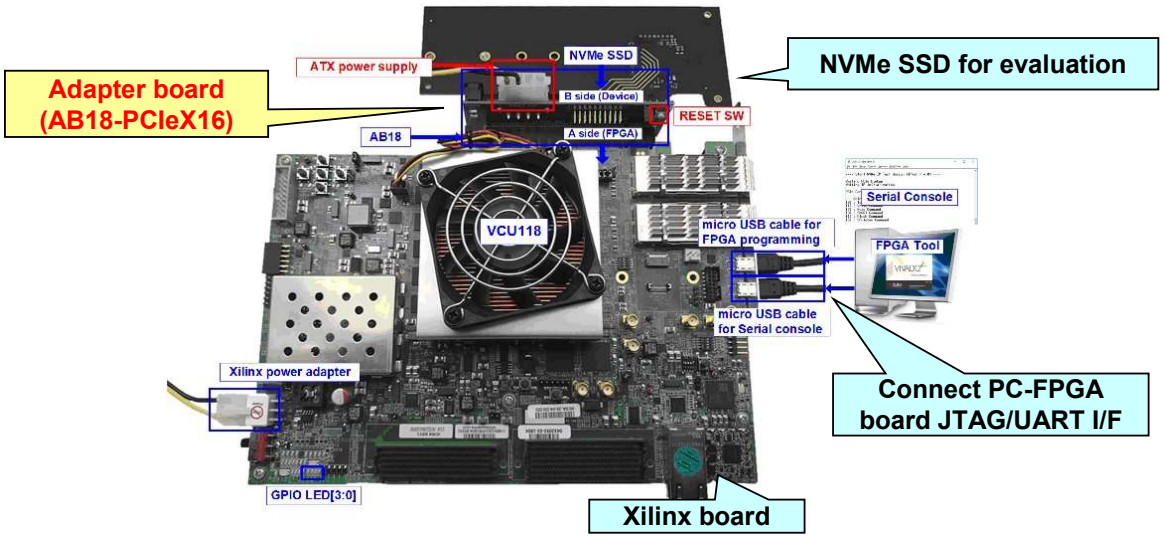
User logic can monitor SSD status such as temperature or total written data count

```
<< SMART Log Information >>
Temperature                    : 32 Degree Celsius
Total Data Read                : 47469 GB
Total Data Written             : 65373 GB
Power On Cycles                : 3991 Times
Power On Hours                 : 79 Hours
Unsafe Shutdowns               : 220 Times
```

SMART command result example

Merit5: Environment

- Real operation check with Xilinx evaluation board
- Free bit-file for evaluation before IP-core purchase



NVMe-IP evaluation environment (Virtex-UltraScale+ evaluation example)

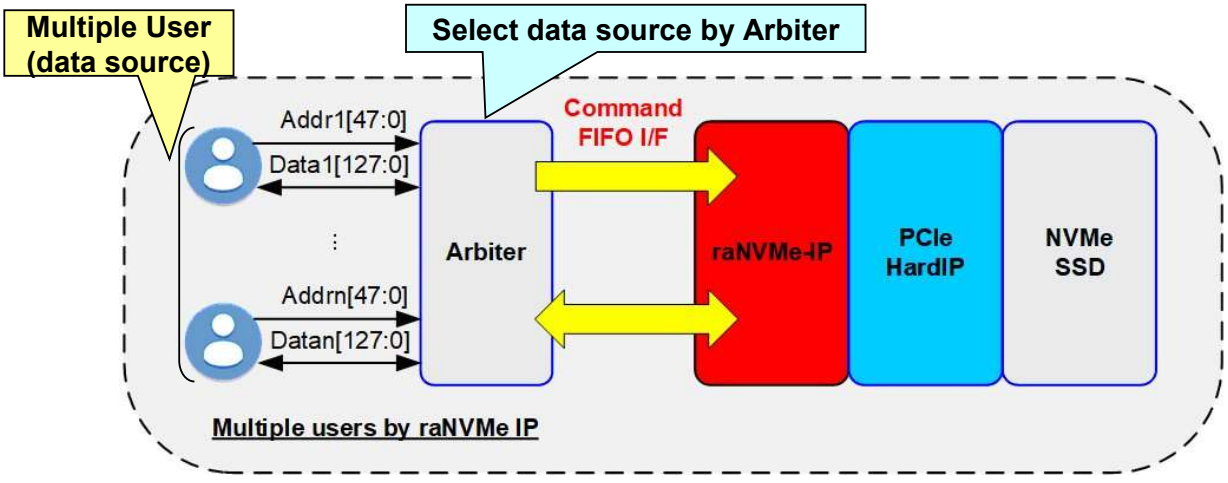
Merit5: Reference Design

- Vivado project attached with raNVMe-IP deliverables
- Full source code (VHDL) except IP core
- Can save user system development duration
  - Confirm real board operation by original reference design.
  - Then modify a little to approach final user product.
  - Check real operation in each modification step.

Short-term development is possible without big turn back

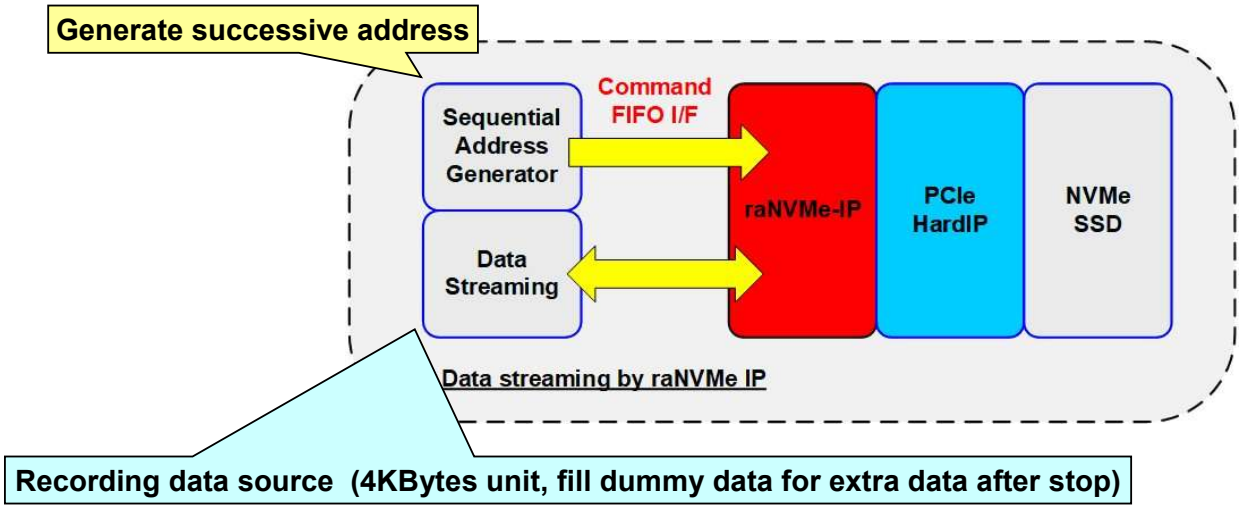
raNVMe-IP Application example1

- Multiple user (host) sharing single SSD system
  - Multiple user logic (sensor data etc.) shares single SSD.
  - Insert Arbiter between multiple user and raNVMe-IP core



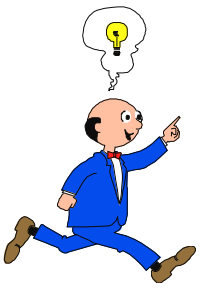
raNVMe-IP Application example2

- **Data Stream Application (Video recording Start-Stop operation)**
  - **Suitable such application that total recording volume is unpredictable.**
  - **Recording by 4KBytes unit, fill dummy data (0-fill) for the last data.**



For more detail

- **Detailed technical information available on the web site.**
  - https://dgway.com/raNVMe-IP_X_E.html
- **Contact**
  - **Design Gateway Co,. Ltd.**
  - **sales@design-gateway.com**
  - **FAX: +66-2-261-2290**



Revision History

Rev.	Date	Description
1.0E	21th-Aug-2020	English version initial release